ANÁLISIS DE CONGLOMERADOS

Otra vía de hacer extracción de conocimiento es mediante la identificación de grupos individuos u observaciones que sean similares. Es decir, identificar grupos de sujetos que son similares con respecto aciertas características.

Sea el caso de identificar estratos sociales mediante los individuos que más se parezcan a determinados tramos de ingresos.

En este caso esta técnica combina observaciones en grupos o conglomerados tales que:

a) Cada grupo es homogéneo con respecto a ciertas características.

b) Cada grupo es diferente del otro con respecto a la misma característica. Es decir, las observaciones de un grupo son diferentes de las observaciones de otro grupo.

Es una técnica adicional de reducción de variables.

A diferencia del análisis factorial que identifica variables latentes (factores, scores de componentes), en este caso, esta técnica es empleada para que los conglomerados de observaciones definan el cluster o conglomerado.

Luego el objetivo es agrupar observaciones tal que cada conglomerado de objetos, individuos, observaciones sen lo más parecidos (homogéneo), con respecto a la variable que esté conglomerando.

Este tipo de análisis es descriptivo sin tener pretensiones de poder inferir sobre sus resultados. Carece de bases estadísticas y su cálculo está diseñado fundamentalmente en proximidades a través de distancias.

Lo que ofrece un análisis de conglomerado es:

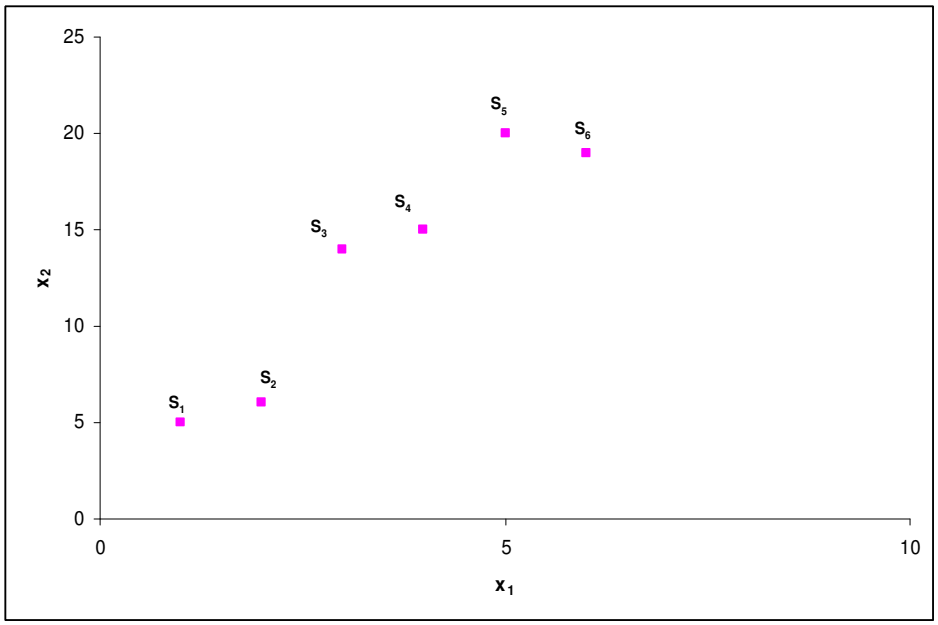
a) A través de la exploración de los datos podemos definir taxonomía sobre los datos.

b) Los datos son reducidos en grupos homogéneos para una misma variable.

c) La definición de los conglomerados ofrece estructuras escondidas representadas en los conglomerados que podrían representar patrones no observables desde las observaciones originales.

Resulta ventajoso para reducir variables categóricas.

Sea por ejemplo un conjunto de individuos u objetos S_1 hasta S_6 tal como se muestra en la figura más abajo. De acuerdo a esta representación, pareciera que se halla agrupado en tres conglomerados. Se sugiere que S_1 y S_2 forme un grupo, S_3 y S_4 forme otro grupo y S_5 y S_6 formarán el tercer grupo.



Tres conglomerados

Más referencias se pueden conseguir en Hair, Anderson, Tatham y Black, que pueden orientar a través de ejemplos de como conducir un análisis exploratorio usando conglomerados.

En este sitio puede ver el desarrollo de la tesis en Tesis (VN 2009)